

# **Collaboration for Extreme Scale Science**

## **Draft Minutes of the Organizing Committee Meeting, October 28, 2011**

### **Present**

Lothar Bauerdick  
Michael Ernst  
Richard Brower  
Dean Williams  
David Skinner  
Mike Papka  
Jeff Kantor  
Ian Foster  
Terence Critchlow  
Miron Livny  
Eli Dart  
Dave Bader  
Rich Carlson

### **First Discussion of the Workshop Agenda**

David Skinner presented the first draft agenda. On the question of breakouts or not, the committee was divided. This issue will be revisited closer to the workshop when we have a better understanding of how the topics covered by the workshop might be split into breakouts. We may need to develop two candidate agendas, with and without breakouts. The breakouts, if any, should be along the lines of the topic classification in the topic table and diagram with any improvements we make to these.

The desirability of having input to the workshop that would make the initial case for each of a number of topic areas was raised and generally supported. This is work the OC still needs to do.

There was support for making overnight homework assignments part of the plan.

### **Review of Proposed Participants**

Most OC members had already checked the availability of their suggested attendees. Final input to the David's and Richard's decision on the first set of invitations should be sent on Monday. This list was already long enough to require some selection so any people who would like to prioritize their proposed names were invited to do so. The selection would (try to) aim at a good representation of science fields, facilities and CS topics and would not explicitly disfavor the suggestions of those who had contributed many names.

### **Continued discussion of topics and their relevance to collaboration**

The topic diagram was discussed:

1. Jeff Kantor suggested we use the broad categories of “People”, “Data”, and “Processing”. This idea seemed to resonate.
2. Ian Foster noted that “Data Management” meant “Data Integration” to some people and “Management of massive scale highly distributed data” to others. Perhaps these areas should be clearly separated.  
Also how services would be delivered (centralized, distributed, via the cloud ...) was another dimension to be considered.
3. Miron Livny pointed out that management of storage was a hard problem.

#### Other issues we should address

- How should DOE/SC exploit current and future tools coming from elsewhere?
- How should DOE/SC exploit its own tools that currently serve narrow constituencies?

It seems clear that the best the workshop could do would be to recommend ongoing processes that would ensure that the program was fully aware of relevant current and likely developments in industry and the world. The barriers to full exploitation of tools developed within DOE/SC were noted. Nobody, except perhaps the science programs, had a mandate to support tools developed within DOE/SC. The science programs would only support tools that were essential for them and there was no effective way for the programs to come together and support cross-cutting tools. Part of the problem was the absence of communication processes – perhaps the workshop could make useful recommendations in this area.

#### AOB and Homework

No arduous homework – just send in suggestions for the first round of invitations no later than Monday. Prioritizations of the names suggested by Labs or interest groups will be welcome.

At a later meeting we will setting homework of Importance/cost/relevance-to-collaboration ratings for the list of topics.